Hoppe-Seyler's Z. Physiol. Chem.
Bd. 363, S. 1397–1407, November 1982

Lipophilin (Proteolipid Apoprotein) of Brain White Matter

Purification and Amino Acid Sequence Studies of the Four Tryptophan Fragments

Wilhelm STOFFEL, Heinz HILLEN, Werner SCHRÖDER and Rainer DEUTZMANN

Institut für Physiologische Chemie der Universität zu Köln

(Received 27 August 1982)

Summary: The chemical cleavage of lipophilin (proteolipid apoprotein) from bovine brain white matter with HBr/dimethyl sulfoxide at the tryptophan residues, under conditions adapted to this hydrophobic protein, releases four fragments with approximate molecular masses 14 kDa (Trp I), 6.8 kDa (Trp IV), 5.2 kDa (Trp III) and 2.1 kDa (Trp II). These fragments have been separated and purified by a combination of solvent distribution, molecular sieve chromatography (Bio-Gel P-150) and high-performance liquid chromatography for automated Edman degradation and combined gas-liquid chromatography/mass spectroscopy.

The complete amino acid sequences of Trp II and III and large sequences of Trp I are reported in this communication. The amino acid sequence of Trp IV and the sequences of peptides releasable from lipophilin by proteolytic enzymes (trypsin, thermolysin, subtilisin, chymotrypsin) have been described in previous reports from this laboratory. Despite two small gaps in the complete primary structure of lipophilin from myelin of central nervous system, our sequence data suggest the arrangement of four long hydrophobic sequences (30–40 apolar amino acid residues) within the hydrophobic core of the myelin lipid bilayer, linked by three hydrophilic regions at the aqueous membrane interphase. These features lend lipophilin the properties of a polytopic membrane protein.

Lipophilin (Proteolipid-Apoprotein) der weißen Substanz des Gehirns. Reinigung und Aminosäuresequenz-Untersuchungen der vier Tryptophan-Fragmente

Zusammenfassung: Die chemische Spaltung des Lipophilins (Proteolipid-Apoprotein) aus der weißen Substanz von Rinderhirn an den Tryptophanresten mit HBr/Dimethylsulfoxid, unter an das hydrophobe Protein adaptierten Reaktionsbedingungen, führt zu vier Fragmenten der Molekularmassen 14 kDa (Trp I), 6.8 kDa (Trp IV), 5.2 kDa (Trp III) und 2.1 kDa (Trp II). Die-

---

Enzymes:
Bacillus subtilis neutral proteinase (thermolysin) (EC 3.4.24.4); Chymotrypsin (EC 3.4.21.1); Subtilisin (EC 3.4.21.14); Trypsin (EC 3.4.21.4).

Abbreviations:
HI = Hydrophobicity index; Trp I–IV = polypeptide fragments I to IV obtained by HBr/dimethyl sulfoxide oxidation at tryptophan residues; Trp I-T_1, Trp I-T_2 . . . = tryptic fragments of Trp I . . . ; DITC glass = *p*-phenylenediisothiocyanate-activated glass; t_R = retention time.

0018-4888/82/0363-1397$02.00

se Polypeptide wurden durch eine Kombination von Lösungsmittelverteilung, Biogel P-150-Chromatographie und Hochdruckflüssigkeitschromatographie für den automatischen Edman-Abbau bzw. die kombinierte Gaschromatographie/Massenspektroskopie getrennt und gereinigt.

Die Gesamtsequenz der Tryptophanbruchstücke II und III sowie lange Aminosäuresequenzen des Tryptophanbruchstücks I werden in dieser Arbeit beschrieben. Die vollständige Sequenz des 6.8 kDa großen Trp IV sowie die Sequenzen der durch proteolytische Enzyme (Trypsin, Thermolysin, Subtilisin, Chymotrypsin) abspaltbaren Bruchstücke wurden in den vorangehenden Arbeiten beschrieben. Obwohl noch zwei kleine Sequenzlücken in der vollständigen Primärstruktur des Lipophilins aus dem Myelin des Zentralnervensystems zu schließen sind, legen unsere bisherigen Sequenzdaten eine Integration von vier langen hydrophoben Sequenzen (30–40 Aminosäurereste) in die hydrophobe Region der Myelinlipid-Doppelschicht nahe. Sie sind durch drei hydrophile in der Membran-Interphase befindliche Regionen verknüpft. Diese Besonderheiten der Primärstruktur verleihen dem Lipophilin die Eigenschaften eines polytopen Membranproteins in der Myelinlipid-Doppelschicht.

Key words: Lipophilin (proteolipid apoprotein), tryptophan cleavage, separation and purification of hydrophobic polypeptides, sequence studies, molecular arrangement of lipophilin in myelin membrane.

Lipophilin (proteolipid apoprotein) is one of the major protein components of the myelin membrane. Although procedures for its isolation have been described already in 1951[1], structural studies have not yet led to a deeper insight into the amino acid sequences. The main reason for this is the insolubility of this 28-kDa hydrophobic protein in aqueous solution.

We have resumed the structural analysis, applying chemical and enzymatic cleavage reactions, and succeeded in the rapid and effective separation and purification of hydrophobic and hydrophilic polypeptides by combined molecular sieve chromatographic separation[2,3]. The amino acid sequence of purified polypeptide fragments was determined by automated Edman degradation[4] and combined gaschromatography/mass spectrometry[2,5–7]. Chemical cleavage at three methionine residues yields four fragments of rather different molecular masses: 18–19 kDa, 5 kDa, 2.1 kDa and 0.7 kDa. We have adapted the cleavage with HBr/dimethyl sulfoxide[8] at the tryptophan residues of the water insoluble protein. Four fragments were released which were more suitable for automated Edman degradation than the cyanogen bromide fragments[9–12], because the largest tryptophan fragment had a molecular mass of 14 kDa, the others of 2.1 kDa, 5.2 kDa and 6.8 kDa. Trp IV (6.8 kDa) resembles the 66 amino acid *C*-terminus, the sequence of which we reported earlier[2] and Trp II (2.1 kDa) was found as part of the sequence of a tryptic fragment, described in our last communication[3].

Here, we describe the tryptophan cleavage procedure of the proteolipid apoprotein, the separation and purification of the tryptophan fragments, the results of sequence analyses of these polypeptides and of tryptic digests thereof. A large number of overlapping peptides obtained by partial acid hydrolysis, derivatized for and analysed by combined gas-liquid chromatography/mass spectrometry, are also included. These fully support the automated sequence data.

The complete amino acid sequences of Trp II, Trp III and IV and the largest part of the 14-kDa Trp I fragment can now be aligned. Disregarding the two small gaps in the complete amino acid sequence, the principle of this hydrophobic membrane protein becomes obvious from the sequence work achieved so far: four hydrophobic amino acid sequences are interrupted by three hydrophilic sequences and bordered by small polar *N*- and *C*-termini. The hydrophobic sequences are sufficiently long enough (between 30 and 40 amino acid residues) to span the lipid bilayer of the myelin membrane. Consideration for a possible integration into the lipid bilayer are put forward.

Materials and Methods

The isolation and purification of bovine myelin lipophilin have been described before[1,2] and the conditions of high-performance liquid chromatography separation of hydrophobic polypeptides originating from the chemical (HBr/dimethyl sulfoxide and CNBr) or enzymatic cleavage (trypsin, thermolysin, subtilisin and chymotrypsin) in our previous papers[2,3].

Modified cleavage at tryptophan residues

The procedure of Wachter and Werhan[8] was modified due to the solubility properties of lipophilin. 300 mg lipophilin was dissolved in a mixture of 500 μl trifluoroacetic acid, 2 ml 12M HCl and 200 μl dimethyl sulfoxide and stirred for 30 min at room temperature. 2 ml 48% aqueous HBr and 200 μl dimethyl sulfoxide were added and left at room temperature for 30 min. 20 ml water was added, the sample centrifuged in a Christ table top centrifuge at 4000 rpm. The protein free supernatant was removed and discarded and the residue dissolved in 5 ml 90% formic acid. The cleavage was completed over 16 h at 40 °C.

Aliquot portions were taken for *high-performance liquid chromatographic analysis* on Lichrosorb Si100 and Si60 with 90% formic acid as solvent[2,3]. Solvent evaporation under vacuo left a residue behind which was extracted twice with 3 ml 60% formic acid. The combined extracts were concentrated to dryness and the residual polypeptides dissolved in 100% formic acid for preparative high-performance liquid chromatography purification of the smaller Trp fragments: Trp III, 5.2 kDa, t_R = 11.76 min and Trp II, 2.1 kDa, t_R = 13.51 min. Rechromatography on a preparative scale yielded homogenous Trp III (t_R = 11.82 min) and Trp II (t_R = 13.94 min). If necessary, the fragments were rechromatographed on a Lichrosorb Si60 and 50 column. The polypeptides were used for automated Edman degradation.

The material, insoluble in 60% aqueous formic acid, contains the two large Trp fragments (Trp I, 14 kDa and Trp IV, 6.8 kDa). These two larger polypeptides were further separated on a *Bio-Gel P-150 column* (4 × 190 mm) with 90% formic acid.

The Trp I fragment (50 mg) was solubilized with 1M sodium hydroxide solution, immediately diluted with 20 ml water and the pH adjusted to 8.3 with ammonium hydrogen carbonate for *trypsin digestion*. The tryptic fragments were separated on a preparative scale by high-performance liquid chromatography on Lichrosorb Si100 and 60 columns. Two homogenous tryptic fragments of Trp I were obtained.

For *automated Edman degradation* a Beckman sequenator, model 890 C, was used with the 0.1M Quadrol program. The phenylthiohydantoin derivatives of the amino acids were identified by high-performance liquid chromatography with a linear gradient program[13] and in addition by thin-layer chromatography[14]. For high-performance liquid chromatography a Waters system model 6000 A, combined with a Kratos-Schöffel UV-monitoring system, model SF770, and a Hewlett-Packard Integrator, model 3390 A, or Beckman Altex pumps, model 100, programmed by the Beckman Controller, model 420, and a Uvicon variable wavelength UV detector (Kontron) connected with a Hewlett-Packard Integrator, model 3390 A, were used. The phenylthiohydantoin derivatives were routinely delivered automatically to the high-performance liquid chromatography system by a Kontron Sampler MSI 660.

For *gas-liquid chromatography/mass spectrometry*, peptides were generated from purified polypeptides by partial acid hydrolysis and derivatized as described in the previous paper[2].

The *separation and computer-assisted analysis* was performed under the recently described conditions and methods[5–7].

Automated solid phase sequencing was performed with a home-made automated sequencer following the design of Machleidt et al.[15] and the program of degradation steps (ref.[16], personal communication).

Results

Tryptophan cleavage of lipophilin and separation of fragments

Following the general strategy in the elucidation of the primary structure of lipophilin from bovine brain white matter this strongly hydrophobic protein was cleaved chemically at three accessible tryptophan residues into four fragments, tryptophan fragment I (Trp I) with molecular mass approximately 14 kDa, resembling the *N*-terminal end and Trp IV with molecular mass 6.8 kDa, the structure of which we reported in a previous communication[2] representing the *C*-terminal sequence of lipophilin.

Trp I at the *N*-terminus is aligned with Trp III of molecular mass 5.2 kDa and Trp II of molecular mass 2 kDa.

The reaction conditions of Wachter and Werhan[8] had to be modified for this water insoluble protein as described under Materials and Methods. The two small polypeptides Trp II (2 kDa) and Trp III (5.2 kDa) were separated from the two larger fragments by repeated extraction with 60% aqueous formic acid, the latter remaining

completely insoluble together with some incompletely cleaved lipophilin. The extract of Trp II and III was separated into components by high-performance liquid chromatography on Lichrosorb Si100 and 60 with formic acid as solvent as described before[2,3]. The purification included rechromatography on Si50 and Si60. TrpIII eluted with t_R 11.82 min and Trp II with 13.94 min (Fig. 1). The purified polypeptides were used directly for automated Edman degradation. The polypeptide mixture, insoluble in 60% formic acid, was dissolved in 100% formic acid and chromatographed on Bio-Gel P-150 with 90% formic acid as solvent. Fig. 2 indicates that, apart form uncleaved lipophilin, the large Trp I fragment (appr. 14 kDa) incompletely separates from Trp IV (6.8 kDa). Aliquot portions of the fractions were analysed by high-performance liquid chromatography on Si100 and Si60 and the enriched fractions combined for preparative high-performance liquid chromatography separation, as described earlier[2,3]. Fig. 3 presents the high-performance liquid chromatography tracing of Trp I and Trp IV used for further sequence studies.

Analysis of tryptophan fragment I (appr. 14 kDa)

Trp I was treated with performic acid[17] and lyophilised. Automated Edman degradation of an aliquot portion of oxidized Trp I over 31 cycles yielded a sequence identical with the 23 amino acid residues respectively obtained from degradation of complete lipophilin described earlier[2,3]. The sequence of the *N*-terminus of bovine lipophilin shows complete homology to the sequence of lipophilin from rat brain[12]. Our sequence is as follows:

Gly-Leu-Leu-Glu-Cys-Cys-Ala-Arg-Cys-Leu-Val-Gly-
-Ala-Pro-Phe-Ala-Ser-Leu-Val-Ala-Thr-Gly-Leu-X-Phe-
-Phe-Gly-Val-Ala-Leu-Phe-

Another aliquot sample of the lyophilised polypeptide (Trp I) was rapidly dissolved in trifluoroacetic acid and the solvent evaporated. The polypeptide became soluble in 1M sodium hydroxide, which was adjusted with ammonium hydrogen carbonate to pH 8.4 for trypsin cleavage.

Two large polypeptides Trp I-T1, molecular mass 5.5 kDa and Trp I-T2, molecular mass 3.5 kDa, were purified by repeated high-performance liquid chromatography for sequence analysis,

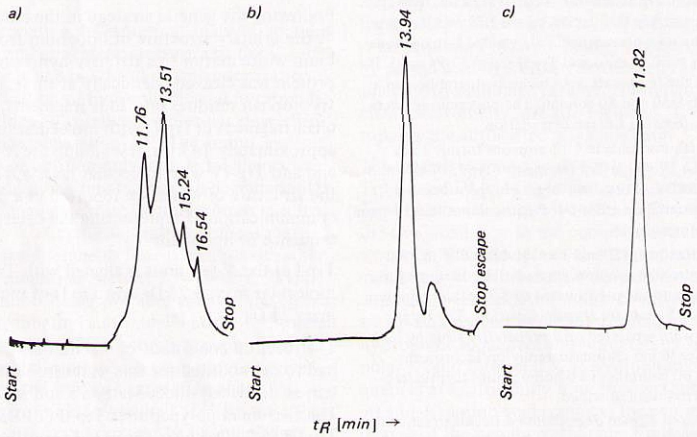

Fig. 1. High-performance liquid chromatography-recording.
a) Total mixture of formic acid (60%) extract, b) Trp II, t_R = 13.94 min, c) Trp III, t_R = 11.82 min after purification.

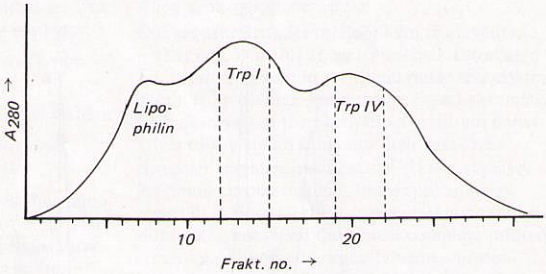

Fig. 2. Recording of Bio-Gel P-150 chromatography (2.5 × 90 cm) of polypeptide mixture insoluble in 60% formic acid, solvent 90% formic acid. Volume of fractions: 5 ml.

Fig. 4. The 5.5-kDa tryptic fragment (T1) of Trp I was sequenced over 10 cycles and yielded the following sequence:

Leu-Ile-Glu-Thr-Tyr-Phe-Ser-Lys-Asn-Tyr-,

whereas the 3.5-kDa tryptic fragment (T2) was washed out after five cycles from which the sequence

Cys-Leu-Val-Gly-Ala-

was derived. Obviously trypsin cleavage had occurred after Arg[8].

Analysis of tryptophan fragment II (appr. 2 kDa)

This polypeptide Trp II was completely analysed by automated Edman degradation. The sequence is as follows:

Leu-Gly-His-Pro-Asp-Lys-Phe-Val-Gly-Ile-Thr-Tyr-Ala--Leu-Thr-Val-Val

Analysis of tryptophan fragment III (appr. 5.2 kDa)

Edman degradation in the liquid sequenator of Trp III turned out to be extremely difficult

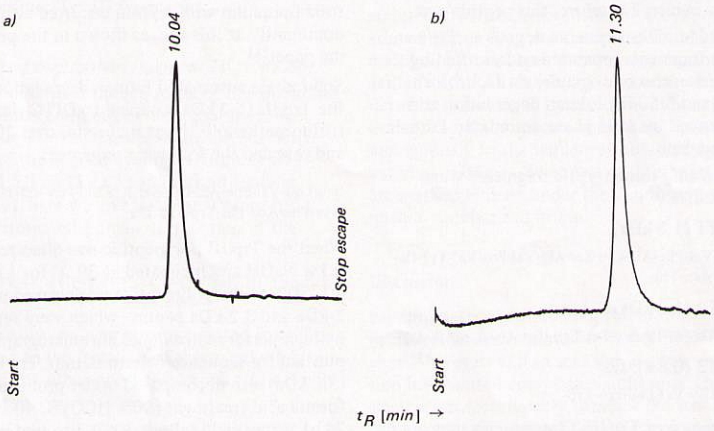

Fig. 3. High-performance liquid chromatography-recording of Trp I (a) and IV (b) purified by preparative high-performance liquid chromatography on Lichrosorb Si100 and Si60.

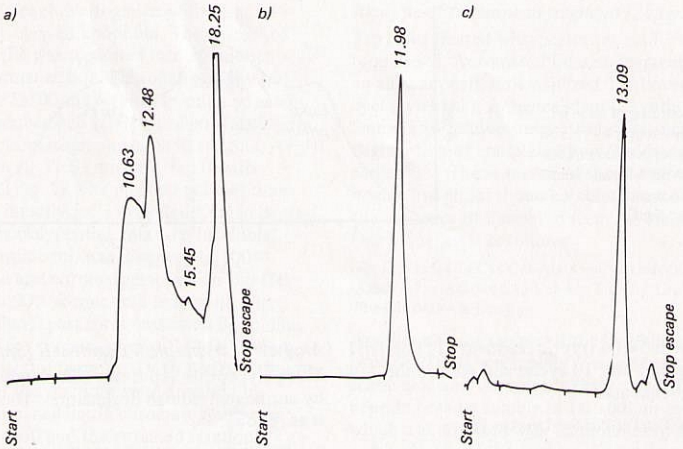

Fig. 4. High-performance liquid chromatography-recording.
a) Mixture of tryptic fragments of Trp I;
b) Trp I-T1, molecular mass 5.5 kDa, t_R = 11.98 min;
c) Trp I-T2, molecular mass 3.5 kDa, t_R = 13.09 min.

because a rapid wash-out effect allowed no more than five cycles. Therefore, this peptide was

a) degraded with trypsin analogous to Trp I and three fragments separated and purified by exclusion chromatography on Lichrosorb Si60 and Si50 for Edman degradation and
b) sequenced by solid phase automated Edman degradation.

The following three tryptic fragments were sequenced:

Trp III-T1 (1.8 kDa):

Leu-Leu-Val-Phe-Ala-Cys-Ser-Ala-Val-Pro-Val-Tyr-Ile-
-Tyr-Phe-Asn-Thr-

Trp III-T2 (1.2 kDa):

Thr-Ser-Ala-Ser-Ile-Gly-Ser-Leu-Cys-Ala-Asp-Ala-Arg-

Trp III-T3 (0.8 kDa):

Met-Tyr-Gly-Val-Leu-Pro-Trp-

The sequence of Trp III-T3 resembles that of the cyanogen bromide fragment II, because of the trypsin hydrolysis at the arginine site preceding methionine. Arginine specific cleavage of total lipophilin with trypsin occurred also predominantly at this site, as shown in the preceding paper[3].

Solid phase automated Edman degradation of the Trp III (5.2 kDa) coupled to DITC-glass in trifluoroethanol[19] was successful over 20 cycles and revealed the following sequence:

Leu-Leu-Val-Phe-Ala-Cys-Ser-Ala-Val-Pro-Val-Tyr-Ile-
-Tyr-Phe-Asn-Thr-Trp-Thr-Thr-

When the Trp III polypeptide was dissolved in 0.1 M NaOH and incubated at 39 °C for 1 h, 40–50% of the polypeptide was cleaved into a 2-kDa and 3.2-kDa peptide which were separated by high-performance liquid chromatography and purified for sequence determination. Trp III (5.2 kDa) also decomposed under prolonged formic acid treatment (90% HCO_2H, 40 °C, 24 h) at low yield (about 20%) into two polypeptides which were separated by high-performance liquid chromatography and sequenced by

liquid phase automated Edman degradation. The 2-kDa polypeptide had the following sequence:

Leu-Leu-Val-Phe-Ala-Cys-Ser-Ala-Val-Pro-Val-Tyr-Ile-
-Tyr-Phe-Asn-Thr-

The first 25 cycles of the 3.2-kDa peptide yielded:

Thr-Thr-Cys-Gln-Ser-Ile-Ala-Ala-Pro-X-Lys-Thr-Ser-Ala-
-Ser-Ile-Gly-Ser-Leu-Cys-Ala-Asp-Ala-Arg-Met-

In the preceding paper[3] we described the amino acid sequence of a tryptic fragment of molecular mass of about 7 kDa. Its sequence overlaps those of the two tryptophan fragments arising from Trp III under the conditions applied here. Furthermore the solid phase sequencing data on Trp III (5.2 kDa) also proved the linear alignment of the two polypeptides.

The resistence of Trp-Thr dipeptides towards tryptophan cleavage has been noticed in an earlier paper[20]. Here we describe another example in which obviously the neighbouring effect of the hydroxy group of threonine on the spirolactone ring leads to a stabilisation of the peptide bond. Trp II and Trp III fragments also resemble the attachment site of one fatty acid residue. Alkaline hydrolysis of the 5.2-kDa Trp III (1M methanolic KOH, 37 °C, 2 h) released 3.4% (w/w) of a mixture of palmitic (41%), stearic (30%) and oleic acid (29%), which is approximately equivalent to one residue/Trp III fragment.

Analysis of tryptophan fragment IV (6.8 kDa)

We have described the sequence of this 66 amino acid residues containing *C*-terminal sequence previously[2]. Whereas the sequence was derived as part of the aligned cyanogen bromide fragments II, III and IV (72 amino-acid residues sequence), here we give additional proof by direct automated Edman degradation of the purified Trp IV over 22 cycles:

Asn-Ala-Phe-Pro-Gly-Lys-Val-Cys-Gly-Ser-Asn-Leu-Leu-
-Ser-Ile-Cys-Lys-Thr-Ala-Glu-Phe-Gln-

A 4-kDa tryptic fragment was isolated from Trp IV similar to the isolation of the tryptic fragments of the other polypeptides over 32 degradation cycles:

Thr-Ala-Glu-Phe-Gln-Met-Thr-Phe-His-Leu-Phe-Ile-Ala-
-Ala-Phe-Val-Gly-Ala-Ala-Ala-Thr-Leu-Val-Ser-Leu-Val-
-Thr-Phe-Met-Ile-Ala-Ala-

Alignment of sequence data

Our sequence studies on lipophilin from white matter of bovine brain, which include automated Edman degradation in the liquid phase sequenator combined with high-performance liquid chromatography-analysis of the phenylthiohydantoin derivatives of the amino acids and their thin-layer chromatographic identification[14] and capillary gas-liquid chromatography/mass spectrometry reported in this and the two previous communications[2,3], unravelled the almost complete primary structure of lipophilin, initially named proteolipid apoprotein[1]. We estimate that about 260 amino acid form this hydrophobic protein, the main protein component of the myelin membrane. Fig. 5 summarizes the sequence data. There are only two small gaps between residues 32 and 44 and around residue 94 according to the tentative numbering in Fig. 5.

Gas-liquid chromatography/mass spectrometry of derivatized peptides released from tryptophan fragments of lipophilin

Aliquot portions of the polypeptides isolated by cleavage at the tryptophan residues of lipophilin and by further trypsin digestion of these tryptophan fragments were digested by partial acid hydrolysis reduced with hexadeuterodiborane and derivatized with *N,N*-diethyltrimethylsilylamine to their respective trimethylsilyl polyaminoalcohols for subsequent gas-liquid chromatography/mass spectrometry as described earlier[2,5,6]. All the di-, tri- and tetrapeptide derivatives identified by the computer-assisted evaluation of the spectra could be assigned unambiguously to the sequences obtained by automated Edman degradation. The peptides are marked by bars under the amino acid sequence summarized in Fig. 5.

Discussion

Lipophilin (proteolipid apoprotein, Folch-Lees-protein) has been isolated for more than 30 years[1], yet its amino acid sequence determination has awaited completion until now. Only small sequences, namely those of the two cyanogen bromide fragments at the *C*-terminus and 18 residues of cyanogen fragment II were sequenced[12] and recently a tryptic fragment (4 kDa) which in-

```
                                                  10                                                  20
Gly-Leu-Leu-Glu-Cys-Cys-Ala-Arg-Cys-Leu-Val-Gly-Ala-Pro-Phe-Ala-Ser-Leu-Val-Ala-
                                                  30                                                  40
-Thr-Gly-Leu-Xaa-Phe-Phe-Gly-Val-Ala-Leu-Phe-----------------------------------
                                                  50                                                  60
------------Lys-Leu-Ile-Glu-Thr-Tyr-Phe-Ser-Lys-Asn-Tyr-Gln-Asp-Tyr-Glu-Tyr-Leu-
            Arg
                                                  70                                                  80
-Ile-Asn-Val-Ile-His-Ala-Phe-Gln-Tyr-Val-Ile-Tyr-Gly-Thr-Ala-Ser-Phe-Phe-Phe-Leu-
                                                  90                                                 100
-Tyr-Gly-Ala-Leu-Leu-Leu-Ala-Xaa-Gly-Phe-Tyr-Thr-------------Lys-----Gly-Leu-Ser-
                                                                Arg
                                                 110                                                 120
-Ala-Thr-Val-Thr-Gly-Gly-Gln-Lys-Gly-Arg-Gly-Ser-Arg-Gly-Gln-His-Gln-Ala-His-Ser-
                                                 130                                                 140
-Leu-Glu-Arg-Val-Cys-His-Cys-Leu-Gly-Xaa-Trp-Leu-Gly-His-Pro-Asp-Lys-Phe-Val-Gly-
                                                 150                                                 160
-Ile-Thr-Tyr-Ala-Leu-Thr-Val-Val-Trp-Leu-Leu-Val-Phe-Ala-Cys-Ser-Ala-Val-Pro-Val-
                                                 170                                                 180
-Tyr-Ile-Tyr-Phe-Asn-Thr-Trp-Thr-Thr-Cys-Gln-Ser-Ile-Ala-Ala-Pro-Xaa-Lys-Thr-Ser-
                                                 190                                                 200
-Ala-Ser-Ile-Gly-Ser-Leu-Cys-Ala-Asp-Ala-Arg-Met-Tyr-Gly-Val-Leu-Pro-Trp-Asn-Ala-
                                                 210                                                 220
-Phe-Pro-Gly-Lys-Val-Cys-Gly-Ser-Asn-Leu-Leu-Ser-Ile-Cys-Lys-Thr-Ala-Glu-Phe-Gln-
                                                 230                                                 240
-Met-Thr-Phe-His-Leu-Phe-Ile-Ala-Ala-Phe-Val-Gly-Ala-Ala-Ala-Thr-Leu-Val-Ser-Leu-
                                                 250                                                 260
-Val-Thr-Phe-Met-Ile-Ala-Ala-Thr-Tyr-Asn-Phe-Ala-Val-Leu-Lys-Leu-Met-Gly-Arg-Gly-

-Thr-Lys-Phe
```

Fig. 5. Proposed primary structure and tentative numbering of lipophilin from bovine white matter. Notice the two sequence gaps beginning at residue 31 and 92.

The bars under sequences of di-, tri- and tetrapeptides refer to peptides identified also by gas-liquid chromatography/mass spectrometry.

cluded 27 amino acids of cyanogen bromide fragment II has been analysed by a combination of solid phase Edman degradation and gas-liquid chromatography/mass spectrometry[18]. The proposed sequence is at variance with that of the segment of the corresponding sequence, which we described in the preceding communication[3] and which is integrated in the almost complete primary structure given in Fig. 5 at three positions: Instead of Val in Pos. 6 of the sequence proposed by Lees et al., which corresponds to Pos. 221 in our tentative numbering, we find Met in repeated

analyses. The amino acid analysis of this polypeptide given by these authors indicates more than two Met. The unknown residue at Pos. 8 (Pos. 224 of Fig. 5) is His and furthermore Val (Pos. 238) is missing in the sequence.

The main obstacle in the elucidation of the primary structure of lipophilin is its hydrophobicity. However, the approach we described before[2,3] and in this paper (solvent distribution, gel exclusion and high-performance liquid chromatography) has also brought this protein in the realm of the regular automated sequencing technique and led to the sequence data summarized above in a rather short time. The aim of these studies on the primary structure of lipophilin is to provide the molecular basis for the observation that this protein may interact with cholesterol and the complex phospho- and sphingolipids of the myelin membrane and thereby contribute to its architecture, which requires a crucial integrity for its function.

Although the intramembranous localization has been questioned from early freeze fracture studies[21], more recent studies clearly visualize the intramembranous protein particles[22]. Intramembranous particles can also be demonstrated when lipophilin has been incorporated into lipid vesicles[23]. The strong interaction with cholesterol and acidic phospholipids has also been demonstrated in artificial lipid monolayers[24]. These spin-labelling technique further indicated hydrophobic interactions[25,26]. The old observation that acidic lipids (phosphatidylserine and phosphatidylinositol) are firmly bound to lipophilin and can be removed only by excessive dialysis against acidic solvents, could be mimicked with bilayer lipid vesicles in these studies. Electrostatic, ionic interactions of 15 phosphatidylserines with 14 basic side chains of lipophilin have been postulated. On the basis of our sequence data the latter figure must be corrected.

A closer look at the primary structure of lipophilin (Fig. 5) with the tentative numbering of amino acid residues, clearly indicates long sequences of uncharged, hydrophobic amino acids. These are linked by shorter hydrophilic sequences with clusters of positively and negatively charged residues. The primary structure of lipophilin is characterized by four hydrophobic segments with more than 30 amino acid residues and one short, 12 residues long sequence with two proline residues. These sequences are continuous with no charged amino acid and sum up to more than 50% of the total polypeptide chain.

Segrest and Feldman[27] described a hydrophobicity diagram of apolar sequences summarizing 774 computer-selected uncharged polypeptides with 10 and more amino acid residues length. Their hydrophobicity scale puts the amino acid residues on a simplified linear scale which is derived from the free energy of transfer of amino acid side chains from an organic to an aqueous environment, described by Nozaki and Tanford[28]. The hydrophobicity index (HI) (sum of hydrophobicity values for the residues of the sequence divided by the number of residues) of the intramembranous domain of glycophorin (23 residues) has been compared with apolar sequences of non-membranous proteins and differs significantly by its high HI of 2.62. The four hydrophobic segments of lipophilin span between residues 10 to very likely 40, residues 59 to 94, with HI of 3.07 for residues 59 to 72 and HI 1.75 for residues 73 to 94. A 17 apolar residues-containing segment between residues 138 to 155 with HI 3.03 and continuing to residue 176 with HI 2.36 represents the most hydrophobic segment. Finally a 36 amino acid residues segment composed of a 13 residues long strongly hydrophobic sequence (residues 219 to 231) with HI = 2.9 and continued by an apolar sequence of 23 amino acids to residue 254 with HI = 1.7 has the properties of membrane-integrated polypeptides. It is quite common to these segments that two and three hydrophobic Phe, Ala, Val and Leu follow in sequence. The hydrophobic segments of about 30 amino acids ordered in an α-helical structure would account for the spanning of the membrane, however folded in β-sheet structure they obtain a length about twice the thickness of the hydrophobic core of the membrane even if we take into account the increased thickness caused by the very long chain fatty acyl residues of the sphingolipids (25–30 Å).

Lipophilin is a basic protein with an isoelectric point above pH 9. On the basis of our present sequence data 16 basic amino acid residues and the N-terminus outnumber the 7 acidic side

chains and the *C*-terminus. They are clustered in rather short hydrophilic segments linking the four hydrophobic segments which span the membrane. One of the charged segments between the first and second hydrophobic segment is negatively charged, whereas all the other hydrophilic segments are positively charged. The tentative arrangement of lipophilin in the myelin membrane is proposed schematically in Fig. 6. We postulate that lipophilin is a polytopic membrane protein[29] with at least four hydrophobic segments spanning or integrated into the membrane core and interacting with cholesterol-rich (44%) bilayer formed from cerebrosides, phosphatidylethanolamines, phosphatidylcholines, sphingomyelins and the alkyl chains of the acidic sulfatides, phosphatidylserines and phosphatidylinositols present in almost equal amounts. It is also worthy to note that, except for two, all the other cysteine residues accumulate in the hydrophilic loops. Disulfide bond formation between the cysteines of the *N*-terminus and those of the hydrophilic loops and between those of the loops themselves could contribute to a compact packing of the intramembranous polypeptide domains. A disulfide bridge between Cys^{214} in the fourth hydrophilic loop with a cysteine residue in the sequence between residue 114 to residue 159 has been demonstrated in the previous paper[3]. Furthermore ionic attractions within two loops of the hydrophilic segments could cluster the hydrophobic intramembranous segments. The negative charges of the second loop Glu^{122} and Asp^{136} and of the third loop (Asp^{189}) and the carboxyl end are neutralized each by an adjacent positively charged amino acid side chain and render these segments strongly basic.

At the moment we can only speculate about the ionic interactions of the basic polypeptide loops

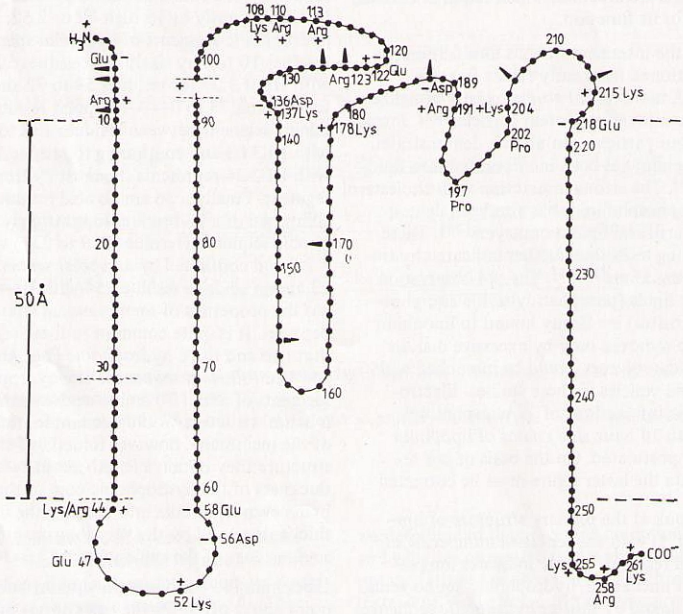

Fig. 6. Schematic presentation of tentatively proposed distribution of hydrophilic and hydrophobic amino acid sequences in the lipid bilayer of myelin membranes. Arrows indicate cysteine residues.

of lipophilin at the membrane interphase, how they contribute together with basic myelin protein to the organisation of the characteristic myelin membrane, about their interactions with acidic polar lipid head groups within the same bilayer or of the opposite membrane and whether they lead to lipid phase separation due to electrostatic forces. However, since lipophilin amounts to about two third of the myelin protein knowledge of the complete primary structure will promote and revitalize research on myelin, particularly with respect to its pathological structural derangements.

Literature

1 Folch-Pi, J. & Lees, M. (1951) *J. Biol. Chem.* **191**, 807–817.
2 Stoffel, W., Hillen, H., Schröder, W. & Deutzmann, R. (1982) *Hoppe-Seyler's Z. Physiol. Chem.* **363**, 855–864.
3 Stoffel, W., Schröder, W., Hillen, H. & Deutzmann, R. (1982) *Hoppe-Seyler's Z. Physiol. Chem.* **363**, 1117–1131.
4 Edman, P. & Begg, G. (1967) *J. Biochem.* **1**, 80–91.
5 Nau, H. & Biemann, K. (1976) *Anal. Biochem.* **73**, 154–174.
6 Nau, H. & Biemann, K. (1976) *Anal. Biochem.* **73**, 175–186.
7 Anderegg, M. (1976) *J. Am. Chem. Soc.* **98**, 3365–3370.
8 Wachter, E. & Werhan, R. (1979) *Anal. Biochem.* **97**, 56–64.
9 Vacher-Leprêtre, M., Nicot, C., Alfsen, A., Jollès, J. & Jollès, P. (1976) *Biochim. Biophys. Acta* **420**, 323–331.
10 Jollès, J., Nussbaum, J.L., Schöntgen, F., Mandel, P. & Jollès, P. (1977) *FEBS Lett.*. **74**, 190–194.
11 Nussbaum, J.L., Ronayrenc, J.F., Mandel, P., Jollès, J. & Jollès, P. (1974) *Biochem. Biophys. Res. Commun.* **57**, 1240–1247.
12 Jollès, J., Schöntgen, F., Jollès, P., Vacher, M., Nicot, C. & Alfsen, A. (1979) *Biochem. Biophys. Res. Commun.* **87**, 619–626.
13 Stoffel, W. & Deutzmann, R., *Hoppe-Seyler's Z. Physiol. Chem.*, in preparation.
14 Bridgen, H. (1975) *FEBS Lett.* **50**, 159–162.
15 Machleidt, W., Hofner, H. & Wachter, E. (1975) in *Solid-Phase Methods in Protein Sequence Analysis* (Laursen, R.A., ed.) pp. 17–30, Pierce Chemical Co., Rockford, Ill.
16 Machleidt, W. & Hofner, H. (1980) in *Methods in Peptide and Protein Sequence Analysis* (Birr, Chr., ed.) pp. 35–47, Elsevier, North-Holland.
17 Hirs, C.H.W. (1956) *J. Biol. Chem.* **219**, 611–621.
18 Lees, M.B., Chao, B.H., Laursen, R.A. & L'Italien, J. (1982) *Biochim. Biophys. Acta* **702**, 117–124.
19 Wachter, E., Machleidt, W., Hofner, H. & Otto, J. (1973) *FEBS Lett.* **35**, 97–102.
20 Ramachandran, L.K. & Witkop, B. (1967) *Methods Enzymol.* **11**, 283–299.
21 Branton, D. (1967) *Exp. Cell Res.* **45**, 703–707.
22 Boggs, J.M. & Moscarello, M.A. (1978) *Biochim. Biophys. Acta* **515**, 1–21.
23 Vail, W.J., Papahadjopoulos, D. & Moscarelleo, M.A. (1974) *Biochim. Biophys. Acta* **345**, 463–467.
24 London, Y., Demel, R.A., Geurts van Kessel, W.S.M., Zahler, P. & van Deenen, L.L.M. (1974) *Biochim. Biophys. Acta* **332**, 69–84.
25 Boggs, J.M., Wood, D.D., Moscarello, M.A. & Papahadjopoulos, D. (1977) *Biochemistry* **16**, 2325–2329.
26 Papahadjopoulos, D., Vail, W.J. & Moscarello, M.A. (1975) *J. Membr. Biol.* **22**, 143–164.
27 Segrest, J. & Feldman, R.J. (1974) *J. Mol. Biol.* **87**, 853–858.
28 Nozaki, Y. & Tanford, C. (1971) *J. Biol. Chem.* **246**, 2211–2217.
29 Lingappa, V. & Blobel, G. (1980) *Recent Prog. Horm. Res.* **36**, 451–475.

Note added in proof (21 October 1982):

After this manuscript had been submitted, we received notice of the paper by Nussbaum J.L., Jollès, J. & Jollès, P. [(1982) *Biochimie* **64**, 405–410] which describes partial sequences of BNPS-skatole fragments of rat brain myelin proteolipid protein (P7). The authors also summarize the present state of the sequence which amounts to a sum of 106 residues in the BNPS-skatole and cyanogen bromide fragments.

Prof. Dr. Dr. Wilhelm Stoffel, Dipl.-Chem. H. Hillen, Dipl.-Chem. W. Schröder and Dr. R. Deutzmann,
Institut für Physiologische Chemie der Universität zu Köln,
Joseph-Stelzmann-Str. 52, D-5000 Köln 41.